SEQUENCE LISTING

(1) GENERAL INFORMATION:

    (i) APPLICANT: Gray, Joe W.
                  Collins, Colin
                  Hwang, Soo-In
                  Godfrey, Tony
                  Kowbel, David
                  Rommens, Johanna

   (ii) TITLE OF INVENTION: Genes From the 20q13 Amplicon and Their
        Uses

  (iii) NUMBER OF SEQUENCES: 55

   (iv) CORRESPONDENCE ADDRESS:
        (A) ADDRESSEE: Townsend and Townsend and Crew LLP
        (B) STREET: Two Embarcadero Center, 8th Floor
        (C) CITY: San Francisco
        (D) STATE: California
        (E) COUNTRY: USA
        (F) ZIP: 94111-3834

    (v) COMPUTER READABLE FORM:
        (A) MEDIUM TYPE: Floppy disk
        (B) COMPUTER: IBM PC compatible
        (C) OPERATING SYSTEM: PC-DOS/MS-DOS
        (D) SOFTWARE: PatentIn Release #1.0, Version #1.30

   (vi) CURRENT APPLICATION DATA:
        (A) APPLICATION NUMBER: US 08/785,532
        (B) FILING DATE: 17-JAN-1997
        (C) CLASSIFICATION:

  (vii) PRIOR APPLICATION DATA:
        (A) APPLICATION NUMBER: US 08/731,499
        (B) FILING DATE: 16-OCT-1996

  (vii) PRIOR APPLICATION DATA:
        (A) APPLICATION NUMBER: US 08/680,395
        (B) FILING DATE: 15-JUL-1996

 (viii) ATTORNEY/AGENT INFORMATION:
        (A) NAME: Quine, Jonathan A.
        (B) REGISTRATION NUMBER: P-41,261
        (C) REFERENCE/DOCKET NUMBER: 023070-068920US

   (ix) TELECOMMUNICATION INFORMATION:
        (A) TELEPHONE: (415) 576-0200
        (B) TELEFAX: (415) 576-0300

(2) INFORMATION FOR SEQ ID NO:1:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 3000 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

   (ii) MOLECULE TYPE: cDNA

(ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..3000
        (D) OTHER INFORMATION: /note= "cDNA clone 3bf4 of 3kb
                                transcript of tyrosine kinase gene A6"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:1:

```
CCGCCGGCCG GGGCGCCTGG CTGCACTCAG CGCCGGAGCC GGGAGCTAGC GGCCGCCGCC        60

ATGTCCCACC AGACCGGCAT CCAAGCAAGT GAAGATGTTA AAGAGATCTT TGCCAGAGCC       120

AGAAATGGAA AGTACAGACT TCTGAAAATA TCTATTGAAA ATGAGCAACT TGTGATTGGA       180

TCATATAGTC AGCCTTCAGA TTCCTGGGAT AAGGATTATG ATTCCTTTGT TTTACCCCTG       240

TTGGAGGACA AACAACCATG CTATATATTA TTCAGGTTAG ATTCTCAGAA TGCCCAGGGA       300

TATGAATGGA TATTCATTGC ATGGTCTCCA GATCATTCTC ATGTTCGTCA AAAAATGTTG       360

TATGCAGCAA CAAGAGCAAC TCTGAAGAAG GAATTTGGAG GTGGCCACAT TAAAGATGAA       420

GTATTTGGAA CAGTAAAGGA AGATGTATCA TTACATGGAT ATAAAAAATA CTTGCTGTCA       480

CAATCTTCCC CTGCCCCACT GACTGCAGCT GAGGAAGAAC TACGACAGAT TAAAATCAAT       540

GAGGTACAGA CTGACGTGGG TGTGGACACT AAGCATCAAA CACTACAAGG AGTAGCATTT       600

CCCATTTCTC GAGAAGCCTT TCAGGCTTTG GAAAAATTGA ATAATAGACA GCTCAACTAT       660

GTGCAGTTGG AAATAGATAT AAAAAAATGAA ATTATAATTT TGGCCAACAC AACAAATACA       720

GAACTGAAAG ATTTGCCAAA GAGGATTCCC AAGGATTCAG CTCGTTACCA TTTCTTTCTG       780

TATAAACATT CCCATGAAGG AGACTATTTA GAGTCCATAG TTTTTATTTA TTCAATGCCT       840

GGATACACAT GCAGTATAAG AGAGCGGATG CTGTATTCTA GCTGCAAGAG CCGTCTGCTA       900

GAAATTGTAG AAAGACAACT ACAAATGGAT GTAATTAGAA AGATCGAGAT AGACAATGGG       960

GATGAGTTGA CTGCAGACTT CCTTTATGAA GAAGTACATC CCAAGCAGCA TGCACACAAG      1020

CAAAGTTTTG CAAAACCAAA AGGTCCTGCA GGAAAAAGAG GAATTCGAAG ACTAATTAGG      1080

GGCCCAGCGG AAACTGAAGC TACTACTGAT TAAAGTCATC ACATTAAACA TTGTAATACT      1140

AGTTTTTTAA AAGTCCAGCT TTTAGTACAG GAGAACTGAA ATCATTCCAT GTTGATATAA      1200

AGTAGGGAAA AAAATTGTAC TTTTTGGAAA ATAGCACTTT TCACTTCTGT GTGTTTTTAA      1260

AATTAATGTT ATAGAAGACT CATGATTTCT ATTTTTGAGT TAAAGCTAGA AAAGGGTTCA      1320

ACATAATGTT TAATTTTGTC ACACTGTTTT CATAGCGTTG ATTCCACACT TCAAATACTT      1380

CTTAAAATTT TATACAGTTG GGCCAGTTCT AGAAAGTCTG ATGTCTCAAA GGGTAAACTT      1440

ACTACTTTCT TGTGGGACAG AAAGACCTTA AAATATTCAT ATTACTTAAT GAATATGTTA      1500

AGGACCAGGC TAGAGTATTT TCTAAGCTGG AAACTTAGTG TGCCTTGGAA AAGCCGCAAG      1560

TTGCTTACTC CGAGTAGCTG TGCTAGCTCT GTCAGACTGT AGGATCATGT CTGCAACTTT      1620

TAGAAATAGT GCTTTATATT GCAGCAGTCT TTTATATTTG ACTTTTTTTT AATAGCATTA      1680

AAATTGCAGA TCAGCTCACT CTGAAACTTT AAGGGTACCA GATATTTTCT ATACTGCAGG      1740

ATTTCTGATG ACATTGAAAG ACTTTAAACA GCCTTAGTAA ATTATCTTTC TAATGCTCTG      1800
```

```
TGAGGCCAAA CATTTATGTT CAGATTGAAA TTTAAATTAA TATCATTCAA AAGGAAACAA      1860

AAAATGTTGA GTTTTAAAAA TCAGGATTGA CTTTTTTCTC CAAAACCATA CATTTATGGG      1920

CAAATTGTGT TCTTTATCAC TTCCGAGCAA ATACTCAGAT TTAAAATTAC TTTAAAGTCC      1980

TGGTACTTAA CAGGCTAACG TAGATAAACA CCTTAATAAT CTCAGTTAAT ACTGTATTTC      2040

AAAACACATT TAACTGTTTT CTAATGCTTT GCATTATCAG TTACAACCTA GAGAGATTTT      2100

GAGCCTCATA TTTCTTTGAT ACTTGAAATA GAGGGAGCTA GAACACTTAA TGTTTAATCT      2160

GTTAAACCTG CTGCAAGAGC CATAACTTTG AGGCATTTTC TAAATGAACT GTGGGGATCC      2220

AGGATTTGTA ATTTCTTGAT CTAAACTTTA TGCTGCATAA ATCACTTATC GGAAATGCAC      2280

ATTTCATAGT GTGAAGCACT CATTTCTAAA CCTTATTATC TAAGGTAATA TATGCACCTT      2340

TCAGAAATTT GTGTTCGAGT AAGTAAAGCA TATTAGAATA ATTGTGGGTT GACAGATTTT      2400

TAAAATAGAA TTTAGAGTAT TTGGGGTTTT GTTTGTTTAC AAATAATCAG ACTATAATAT      2460

TTAAACATGC AAAATAACTG ACAATAATGT TGCACTTGTT TACTAAAGAT ATAAGTTGTT      2520

CCATGGGTGT ACACGTAGAC AGACACACAT ACACCCAAAT TATTGCATTA AGAATCCTGG      2580

AGCAGACCAT AGCTGAAGCT GTTATTTTCA GTCAGGAAGA CTACCTGTCA TGAAGGTATA      2640

AAATAATTTA GAAGTGAATG TTTTTCTGTA CCATCTATGT GCAATTATAC TCTAAATTCC      2700

ACTACACTAC ATTAAAGTAA ATGGACATTC CAGAATATAG ATGTGATTAT AGTCTTAAAC      2760

TAATTATTAT TAAACCAATG ATTGCTGAAA ATCAGTGATG CATTTGTTAT AGAGTATAAC      2820

TCATCGTTTA CAGTATGTTT TAGTTGGCAG TATCATACCT AGATGGTGAA TAACATATTC      2880

CCAGTAAATT TATATAGCAG TGAAGAATTA CATGCCTTCT GGTGGACATT TTATAAGTGC      2940

ATTTTATATC ACAATAAAAA TTTTTTCTCT TTAAAAAAAA AAAACAAGAA AAAAAAAAAA      3000
```

(2) INFORMATION FOR SEQ ID NO:2:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 723 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: cDNA

     (ix) FEATURE:
          (A) NAME/KEY: -
          (B) LOCATION: 1..723
          (D) OTHER INFORMATION: /note= "cDNA clone 1b11 of 3.5kb
                           transcript"

     (xi) SEQUENCE DESCRIPTION: SEQ ID NO:2:

```
TGGAAGCTGT CATGGTTACC GTCTCTAACG TTGGACTCTT AAGAAAATGA TTATTCCTGG        60

TTTCTAGACA GGCCAAATGT AATTCACCTA CGTGGCAGAT TAAAGAGGTG GGCTTACTAG       120

ATTTGATTGG GTATTGAGCA TGCTCTGAAT GACAGTCCCC AAAAAGGACC TCTTATCCGT       180
```

```
TCTTCCCCTT GGGGAAGGGC TTTTGCCACT TCCATGTCAA TGTGGCAGTT GAGCTTGGAA        240

ATTGGTGCGT TGTACAACAT AAGCATTACT TCTCCAAGAT GTGCCTGTGT AGAAATGGTC        300

ATAGATTCAA AACTGTAGCT ACTATGTGGA CAGGGGGGCA GCAAGGACCC CACTTTGTAA        360

AACATGTTTT GGGGGAATGT TTTGTTTTTC ATTTTCTTAT TACCTGGCAA AATAATCCAG        420

GTGGTGTGTG AGTCACCAGT AGAGATTATA AAGTCCAAGG AAGTAGAATC AGCCTTACAA        480

ACAGTGGACC TCAACGAAGG AGATGCTGCA CCTGAACCCA CWGAAGCGAA ACTCAAAAGA        540

GAAGAAAGCA AACCAAGAAC CTCTCTGATG RCGTTTCTCA GACAAATGGT AAGCCCCTTA        600

CTTCCAGTAT AGGAAACCTA AGATACCTAG AGCGGCTTTT GGGAACAATG GGCTCATGCC        660

ACAGGTAGTA GGAGACATAA TTGTAGCTGG TGTGTATGGA ATGTGAATGG AATATGGATT        720

GCG                                                                     723
```

    (2) INFORMATION FOR SEQ ID NO:3:

            (i) SEQUENCE CHARACTERISTICS:
                   (A) LENGTH: 1507 base pairs
                   (B) TYPE: nucleic acid
                   (C) STRANDEDNESS: single
                   (D) TOPOLOGY: linear

           (ii) MOLECULE TYPE: cDNA

           (ix) FEATURE:
                   (A) NAME/KEY: -
                   (B) LOCATION: 1..1507
                   (D) OTHER INFORMATION: /note= "cDNA clone cc49 of 6-7kb
                                          transcript with homology to C2H2 zinc
                                          finger genes"

           (xi) SEQUENCE DESCRIPTION: SEQ ID NO:3:

```
GCAGGTTGCT GGGATTGACT TCTTGCTCAA TTGAAACACT CATTCAATGG AGACAAAGAG         60

CACTAATGCT TTGTGCTGAT TCATATTTGA ATCGAGGCAT TGGGAACCCT GTATGCCTTG        120

TTTGTGGAAA GAACCAGTGA CACCATCACT GAGCTTCCTA AAAGTTCGAA GAAGTTAGAG        180

GACTATACAC TTTCTTTTGA ACTTTTATAA TAAATATTTG CTCTGGTTTT GGAACCCAGG        240

ACTGTTAGAG GGTGAGTGAC AGGTCTTACA GTGGCCTTAA TCCAACTCCA GAAATTGCCC        300

AACGGAACTT TGAGATTATA TGCAATCGAA AGTGACAGGA AACATGCCAA CTCAATCCCT        360

CTTAATGTAC ATGGATGGCC AAGAGTGATT GGCAGCTCTC TTGCCAGTCC GATGGAGATG        420

GAGATGCCTT GTCAATGAAA GGGCCCNCTG TTGTCAATTC CGAGCTACAC AAAGAAAAAA        480

ATGTCAATCC GAATCGAGGG GAATATGCCC TTGGATTGCA TGTTCTGCAG CCAGACCTTC        540

ACACATTCAG AAGACCTTAA TAAACATGTC TTAATGCAAC ACCGGCCTAC CCTCTGTGAA        600

CCAGCAGTTC TTCGGGTTGA AGCAGAGTAT CTCAGTCCGC TTGATAAAAG TCAAGTGCGA        660

ACAGAACCTC CCAAGGAAAA GAATTGCAAG GAAAATGAAT TTAGCTGTGA GGTATGTGGG        720

CAGACATTTA GAGTCGCTTT TGATGTTGAG ATCCACATGA GAACACACAA AGATTCTTTC        780
```

```
ACTTACGGGT GTAACATGTG CGGAAGAAGA TTCAAGGAGC CTTGGTTTCT TAAAAATCAC      840

ATGCGGACRC ATAATGGCAA ATCGGGGGCC AGAAGCAAAC TGCAGCAAGG CTTGGAGAGT      900

AGTCCAGCAA CGATCAACGA GGTCGTCCAG GTGCACGCGG CCGAGAGCAT CTCCTCTCCT      960

TGCAAAATCT GCATGGTTTG TGGCTTCCTA TTTCCAAATA AAGAAAGTCT AATTGAGCAC      1020

CGCAAGGTGC ACACCAAAAA AACTGCTTTC GGTACCAGCA GCGCGCAGAC AGACTCTCCA      1080

CAAGGAGGAA TGCCGTCCTC GAGGGAGGAC TTCCTGCAGT TGTTCAACTT GAGACCAAAA      1140

TCTCACCCTG AAACGGGGAA GAAGCCTGTC AGATGCATCC CTCAGCTCGA TCCGTTCACC      1200

ACCTTCCAGG CTTGGCAKCT GGCTACCAAA GGAAWAGTTG CCATTTGCCA AGAAGTGAAG      1260

GAATTGGGGC AAGAAGGGAG CACCGACAAC GACGATTCGA GTTCCGAGAA GGAGCTTGGA      1320

GAAACAAATA AGAACCATTG TGCAGGCCTC TCGCAAGAGA AAGAGAAGTG CAAACACTCC      1380

CACGGCGAAG CGCCCTCCGT GGACGCGGAT CCCAAGTTAC CCAGTAGCAA GGAGAAGCCC      1440

ACTCACTGCT CCGAGTGCGG CAAAGCTTTC AGAACCTACC ACCAGCTGGT CTTGCACTCC      1500

AGGGTCC                                                               1507
```

(2) INFORMATION FOR SEQ ID NO:4:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 2605 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..2605
        (D) OTHER INFORMATION: /note= "cDNA clone cc43 of 4 kb
            transcript"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:4:

```
CAAGCTCGAA ATTAACCCTC ACTAAAGGGA ACAAAAGCTG GAGCTCCACC GCGGTGGCGG      60

CCGCTCTAGA ACTAGTGGAT CCCCCGGGCT GCAGGAATTC GGCACGAGCT GGGCTACTAC      120

GATGGCGATG AGTTTCGAGT GGCCGTGGCA GTATCGCTTC CCACCCTTCT TTACGTTACA      180

ACCGAATGTG GACACTCGGC AGAAGCAGCT GGCCGCCTGG TGCTCGCTGG TCCTGTCCTT      240

CTGCCGCCTG CACAAACAGT CCAGCATGAC GGTGATGGAA GCTCAGGAGA GCCCGCTCTT      300

CAACAACGTC AAGCTACAGC GAAAGCTTCC TGTGGAGTCG ATCCAGATTG TATTAGAGGA      360

ACTGAGGAAG AAAGGGAACC TCGAGTGGTT GGATAAGAGC AAGTCCAGCT TCCTGATCAT      420

GTGGCGGAGG CCAGAAGAAT GGGGGAAACT CATCTATCAG TGGGTTTCCA GGAGTGGCCA      480

GAACAACTCC GTCTTTACCC TGTATGAACT GACTAATGGG AAGACACAG AGGATGAGGA      540

GTTCCACGGG CTGGATGAAG CCACTCTACT GCGGGCTCTG CAGGCCCTAC AGCAGGAGCA      600
```

```
CAAGGCCGAG ATCATCACTG TCAGCGATGG CCGAGGCGTC AAGTTCTTCT AGCAGGGACC        660

TGTCTCCCTT TACTTCTTAC CTCCCACCTT TCCAGGGCTT TCAAAAGGAG ACAGACCCAG        720

TGTCCCCCAA AGACTGGATC TGTGACTCCA CCAGACTCAA AAGGACTCCA GTCCTGAAGG        780

CTGGGACCTG GGGATGGGTT TCTCACACCC CATATGTCTG TCCCTTGGAT AGGGTGAGGC        840

TGAAGCACCA GGGAGAAAAT ATGTGCTTCT TCTCGCCCTA CCTCCTTTCC CATCCTAGAC        900

TGTCCTTGAG CCAGGGTCTG TAAACCTGAC ACTTTATATG TGTTCACACA TGTAAGTACA        960

TACACACATG CGCCTGCAGC ACATGCTTCT GTCTCCTCCT CCTCCCACCC CTTTAGCTGC       1020

TGTTGCCTCC CTTCTCAGGC TGGTGCTGGA TCCTTCCTAG GGGATGGGGG AAGCCCTGGC       1080

TGCAGGCAGC CTTCCAGGCA ATATGAAGAT AGGAGGCCCA CGGGCCTGGC AGTGAGAGGT       1140

GTGGCCCCAC ACCGATTTAT GATATTAAAA TCTCAACTCC CAAAAAAAAA AAAAAAAAA       1200

CTGAGACTAG TTCTCTCTCT CTCGAGAACT AGTCTCGAGT TTTTTTTTT TTTTTTTTTT       1260

TTTTTTTTTT TTTTTTTTTG GCTTTAAGGA TTTATTTATT GTTTCCTCTT TACAGTGTCC       1320

ACTTTTCTCT ACTTAATACT ACTTTCCAGT CTCAGAAGCC CAGAGGGAAA AAAAAAAGAC       1380

CATGAATCTT CCTCTCCCAG ATTAAAGTAC ACACTTTGGA AAACAGATTG GAAAACCTTT       1440

CTGAAAAAAG TTGACTGAAA CTCCAAACCA ACATGCCATA TTGTTGATGT TGCTCATGAA       1500

AATTGTTAAA AACCTGTTCT AGATAAAGAA CAGTCTCAAG TTTTTGTACA GCCTACACAT       1560

AGTACAAGGG TCCCCTATGA TGATTCTTCT GTAGGACGAA ATAATGTAAT TTTTTCAGTT       1620

TCTGGTTTAT AACTCTCTCG ATCTCAGAGT TGACTGATTA AAACACCTAC TCATGCAACA       1680

GAGAATAAAG CACTCATATT TTTATAAATT ATATGGACCA AACTATTTTG GAAATCTTAT       1740

CTATTGGAGA CACAATATGC TGGACTAAAG CAATAATTAT TTTATTCTCA ATGTCTGTGC       1800

TAACCTCAAT GACTTAGAAT GCTTTGCTAT ATTTTGCCTC TATGCCTCAA CCACACTGGC       1860

TTTCTTTTAG CTCTTGAACA AGCCAAACTG CTTCCTGCCT CAGGACCAGA TATTTTGGGA       1920

CTTCTCTTAA GAATTCTATT TCCTTAATTC TTTATCTGGG TAACTTAGTT TTATCCAACA       1980

CTTCAGATCC TGCCGTAAAA ACTCTTCTTA TAGAAGCCTG TCATGACACT GTCTCTCTTC       2040

TCCAACATAC TCACCAGCAC ACATGTAGAC TAGATTAGAA CCTCCTGTTT TTCTTTTTCA       2100

TACTTTTCTC TATCATGCTT CCCTCCATTA TAATATTTTT ATTATGTGTG TGAATGTCTG       2160

CCCCAAGTCA GTTTCCTCAC TAAACTATAA ACTCCGTAAA GCTGGGATCC TTCCAATTTT       2220

GATCACCACT TAGTACAGTA GGAACACAGT AAAGATTCAA TTGGTATTTG TGGAATGAAT       2280

GAATGAATTG TTTTGCTAGT AAAGTCTGGG GGAACCCAGG TGAGAAGAGC CTAGAAAGCA       2340

GGTCGAATCC AAGGCTAGAT AGACTTAGTG TTACTCAAGA AAGGGTAGCC TGAAAATAAA       2400

GGTTCAAATT ATAGTCAAGA ATAGTCAAGA CATGGGCAAG ACAAGAGTGC TGCTCGTGCC       2460

GAATTCGATA TCAAGCTTAT CGATACCGTC GACCTCGAGG GGGGGCCCGG TACCCAATTC       2520

GCCCTATAGT GAGTCGTATT ACAATTCACT GGCCGTCGTT TTACAACGTC GTGACTGGGA       2580

AAACCCTGGC GTTACCCAAC TTAAT                                           2605
```

(2) INFORMATION FOR SEQ ID NO:5:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1288 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..1288
        (D) OTHER INFORMATION: /note= "cDNA clone 41.1 with homology
                             to homeobox T shirt gene from
                             Drosophila"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:5:

```
GAGGGCAGCG AGAAGGAGAA ACCCCAGCCC CTGGAGCCCA CATCTGCTCT GAGCAATGGG        60

TGCGCCCTCG CCAACCACGC CCCGGCCCTG CCATGCATCA ACCCACTCAG CGCCCTGCAG       120

TCCGTCCTGA ACAATCACTT GGGCAAAGCC ACGGAGCCCT TGCGCTCACC TTCCTGCTCC       180

AGCCCAAGTT CAAGCACAAT TTCCATGTTC CACAAGTCGA ATCTCAATGT CATGGACAAG       240

CCGGTCTTGA GTCCTGCCTC CACAAGGTCA GCCAGCGTGT CCAGGCGCTA CCTGTTTGAG       300

AACAGCGATC AGCCCATTGA CCTGACCAAG TCCAAAAGCA AGAAAGCCGA GTCCTCGCAA       360

GCACAATCTT GTATGTCCCC ACCTCAGAAG CACGCTCTGT CTGACATCGC CGACATGGTC       420

AAAGTCCTCC CCAAAGCCAC CACCCCAAAG CCAGCCTCCT CCTCCAGGGT CCCCCCCATG       480

AAGCTGGAAA TGGATGTCAG GCGCTTTGAG GATGTCTCCA GTGAAGTCTC AACTTTGCAT       540

AAAAGAAAAG GCCGGCAGTC CAACTGGAAT CCTCAGCATC TTCTGATTCT ACAAGCCCAG       600

TTTGCCTCGA GCCTCTTCCA GACATCAGAG GGCAAATACC TGCTGTCTGA TCTGGGCCCA       660

CAAGAGCGTA TGCAAATCTC TAAGTTTACG GGACTCTCAA TGACCACTAT CAGTCACTGG       720

CTGGCCAACG TCAAGTACCA GCTTAGGAAA ACGGGCGGGA CAAAATTTCT GAAAAACATG       780

GACAAAGGCC ACCCCATCTT TTATTGCAGT GACTGTGCCT CCCAGTTCAG AACCCCTTCT       840

ACCTACATCA GTCACTTAGA ATCTCACCTG GGTTTCCAAA TGAAGGACAT GACCCGCTTG       900

TCAGTGGACC AGCAAAGCAA GGTGGAGCAA GAGATCTCCC GGGTATCGTC GGCTCAGAGG       960

TCTCCAGAAA CAATAGCTGC CGAAGAGGAC ACAGACTCTA AATTCAAGTG TAAGTTGTGC      1020

TGTCGGACAT TTGTGAGCAA ACATGCGGTA AAACTCCACC TAAGCAAAAC GCACAGCAAG      1080

TCACCCGAAC ACCATTCACA GTTTGTAACA GACGTGGATG AAGAATAGCT CTGCAGGACG      1140

AATGCCTTAG TTTCCACTTT CCAGCCTGGA TCCCCTCACA CTGAACCCTT CTTCGTTGCA      1200

CCATCCTGCT TCTGACATTG AACTCATTGA ACTCCTCCTG ACACCCTGGC TCTGAGAAGA      1260

CTGCCAAAAA AAAAAAAAAA AAAAATTC                                        1288
```

(2) INFORMATION FOR SEQ ID NO:6:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 2821 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..2821
        (D) OTHER INFORMATION: /note= "cDNA clone GCAP encodes a
                             guanino cyclase activating protein
                             (GCAP)"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:6:

```
ATCCTAAGAC GCACAGCCTG GGAAGCCAGC ACTGGGGAAG TGGTGCTGAG GGATGTGGGT      60

CACTGGGGTG AAGGTGGAGC TTTCAGGGTC TCCCGTCAAT GCAGCTGAGT TTTCTTTGGC     120

AGGGAATTTA CCAGCTGAAG AAAGCCTGCC GGCGAGAGCT ACAAACTGAG CAAGGCCAGC     180

TGCTCACACC CGAGGAGGTC GTGGACAGGA TCTTCCTCCT GGTGGATGAG AATGGAGATG     240

GTAAGAGGGG CAGAGATGGG GAGAGTGCTG TCCACTCTGC ATCATCGCCA CTTTCTGGCC     300

GCACGTCCTT GGGCAAGGCC CTCCACCTTC CAACCCTGGG GTCCTCATCT GTGAGAAGGC     360

TGTGGAGAAG ATGTCATGAA CTAACAAAGG GACTCATGAG CACGTGTTTG TAGGAGTGAC     420

TAAAAGTCCT ACAGGAGTTG CTGATGGAGG CCAGGCACGC AGAATAGAAA GAATAGGAAC     480

TTTGGAGTCA GGCAGGGAGT GATATATTGA GCTTCTCGTC CTAGTCTCAA TTTCCTCATC     540

TGGAAAATGG GGATAATAAT AGTGGTTGAG AGGAATGAAT AGGATAATGT GTTTAAGAGC     600

AGGCATAGGG TAGACCTCCA TTCAGGCTGC TTGGGCTTTC CTCCCTGTAG CCCAAAGCCC     660

AGCCTCAGGG CTATGTGGGG AGAGAGCTGG CTTGGAATAC ACACTTGAGC CCTCCAGCTC     720

TCTCAGCTCC ACCCAGCATT TCCGTGGTAC CATGCGCAAA AGTAAAACTT CAATTCATCA     780

GCAAAGAAAG CCCCTTAAAG GTGGCAGGAG ACTCCTGGAG ATTCAGACAC CTGACAAGCC     840

GCAAGCTTGA GGTCTGAGAC TGCAGGATAG TTGGCATAAG ACGTGTAGGC GCATCCTGGG     900

AGCGAGGTCT CTCCTCCTGC CCCCAGACCC AGGTCTCCCC TTCTTCTACA TGACCACCTC     960

TCCTCCCCCT TGCTCAGGCC AGCTGTCTCT GAACGAGTTT GTTGAAGGTG CCCGTCGGGA    1020

CAAGTGGGTG ATGAAGATGC TGCAGATGGA CATGAATCCC AGCAGCTGGC TCGCTCAGCA    1080

GAGACGGAAA AGTGCCATGT TCTGAGGAGT CTGGGGCCCC TCCACGACTC CAGGCTCACC    1140

CAGGTTTCCA GGGTAGTAGG AGGGTCCCCT GGCTCAGCCT GCTCATGCCC ACTCTTCCCC    1200

TGGTGTTGAC TTCCTGGCAC CCCCTGTGCA GGGCTGAGTG GGGATGGGGA AGGGCTGCTG    1260

GGTTTGAAGT GGCCAACAGG GCATAGTCCA TTTTGGAGGA GTCCCTGGGA TGGTGAAGGG    1320

AATTCAGTTA CTTTTCCTGT TCAGCCGCTC CTGGGAGGAC TGTGCCTTGG CTGGGTGGTT    1380
```

```
GTGGGGCTCC CACAGTTTCT GGGTGTTCTC AGTTGGAAGC AAGAGCCAAC TGAGGGGTGA      1440

GGGTCCCACA GACCAAATCA GAAATGAGAA CACAAAGACT GGTAGGAGGC AGGGGTGGGA      1500

GGGTGTTGAG ACTGAAGAAA AGGCAGGAGT TGCCGGGCAC GGTGGCTCAC GCCTGTAATC      1560

CCAGCACTTT GGGAGGCCGA GGCGGGCAGA TCACGAGGTC AGGAGATCGA GACCATCCTG      1620

GCTAACACGG GGTGAAACCC CGTCTCTACT AAAAATACAA AAAATCAGCC GGGTGAGGTG      1680

GCGGGCGCCT GTAGTCCCAG CTACTCAGGA GGCTGAGGCA AGAGAATGGC GTGAACCCCA      1740

GGGGGCCGAG CCTACAGTGA GCCGAGATTG CGCCACTGCA CTCCAGCCTG GACGACAGTG      1800

AGACTCCGTC TCAAAAAAAA AAAAGAAAG AAAAGAAAG GCAGGAGTTT TGGGGGGCAG      1860

GGGGCAGCAA TAATTCTATA ACTTCCGGGA TGCTGAGGGG CGTTCATGGG GAGGACCCTG      1920

GCCTCCTCCT CCCCAAGGCA TCCTCACCAG TGGTGTCAAC AGGAAAAATG GCAGCAAATA      1980

CGCTGCAGGC TGTGGTCTTT CTGCCTTTGA AAGGGTCAGC TGTACTTAAA GGGACTGTTT      2040

CAGCTCTGCC TGGGTGCTGC TCTGGGACCC CCTGCTGCCA ACCCACCACT CCCCCAACAA      2100

TCCTCTCTTT CCATCCATAT CCCCCAGTAT GGACCTTCCA CAACTCCCAG CCATAAGCTG      2160

AATGTTTCTC TTTAAAGGAT GGAGAAAACT TCTGTCTGTC TCTGGCAAGA ATTGGGGGAC      2220

TGTTGACTGG GATTGTGGGC TGGGCTTGGC TTCTAACTGC TGTGTGACCC AAGACAGCCA      2280

CTTCTCCTCC CTAACCTTGG TTATGTCTTG GCAGCACAGT GAGCAGGTCG GACTAGGCGA      2340

ACAGTTTTGG ATTATTGTGT TTTTAGATGT GGAATTATTT TTTGTTATAT AAACTCTTAT      2400

GTGTAACCCC AATATAGAAA CTAGATTAAA AGGGAGTCTC TCTGGTTGAA AGGGGAGCTG      2460

AGTACCCTCT GGAACTGGAG GCACCTCTGA AAAAAGCAAA CTGAAAACCA GTGCCCTGGG      2520

TCACTGTTAC TCCTATAAGA CAGTTTAAAG TGAGACCTGG AAAAACATTT GCTTTACCTT      2580

GAATAGATAG GTTTTTATGT TGGTATATAA GAAATAAAAC TAACCTATTA ACCCTGAGAC      2640

TTTACAGGTG TGTTATTTCA TATGATAGTC ATATAAAATT TCCTTTAGAC ATCAATTTTA      2700

GGTAAAAAAT AATTGATTAG AAAAATATTG GCCAGGTGCA GCAGCTCACA CCTGCAATCC      2760

CAGGACTTTG GGAGGCCGAG GCGGGTGGAT CACCTGAGGT CAGGGGTTCA AGACCAGCCT      2820

G                                                                     2821
```

(2) INFORMATION FOR SEQ ID NO:7:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1205 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..1205
        (D) OTHER INFORMATION: /note= "cDNA clone 1b4 for a serine
                            threonine kinase"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:7:

```
GCGCGCGTGA GTCCGCCCCC CCAGTCACGT GACCGCTGAC TCGGGGCGTT CTCCACTATC      60

GCTTACCTAC CTCCCTCTGC AGGAACCCGG CGATATGGCT GCCGCTGTGC CCCGCGCCGC     120

ATTTCTCTCC CCGCTGCTTC CCTTCTCCTG GGCTTCCTGC TCCTCTCCGC TCCGCATGGC     180

GGCAGCGGCC TGCACACCAA GGCGCCCTTC CCCTGGATAC GGTCACTTTC TACAAGGTCA     240

TTCCCAAAAG CAAGTTCGTC TGGTGAAGTT CGACACCCAG TACCCCTACG GTGAGAAGCA     300

GGATGAGTTC AAGCGTCTTC TGAAAACTCG GCTTCCAGCG ATGATCTCTT GGTGGCAGAG     360

GTGGGGATCT CAGATTATGT GACAAGCTGA ACATGGAGCT GAGTGAGAAA TACAAGCTGG     420

ACAAAGAGAG CTACCCATCT TCTACCTCTT CCGGGATGGG GACTTTGAGA ACCCAGTCCC     480

ATACACTGGG GCAGTTAGGT TGGAGCCATC CAGCGCTGGC TGAAGGGGCA AGGGGTCTAC     540

CTAGGTATGC CTGGTGCCTG CCTGTATACG ACGCCCTGGC CGGGGAGTTC ATCAGGGCCT     600

CTGGTGTGGA GGCCGCCAGG CCCTCTTGAA GCAGGGGCAA GATAACCTCT CAAGTGTGAA     660

GGAGACTCAG AAGAGTGGGC CGAGCAATAC CTGAAGATCA TGGGGAAGAT CTTAGACCAA     720

GGGGAGCACT TCCAGCATCA GAGATGACAC GGATCGCCAG GCTGATTGAG AAGAACAAGA     780

TGAGTGACGG CAGAAGGAGG AGCTCCAGAA GAGCTTAAAC ATCCTGACTG CCTTCCAGAA     840

GAAGGGGGCC GAGAAAGAGG AGCTGTAAAA AGGCTGTCTG TGATTTTCCA GGGTTTGGTG     900

GGGGTAGGGA GGGGANAGTT AACCTGCTGG CTGTGANTCC CTTGTGGAAT ATAAGGGGGY     960

MSKGGGAAAA GWGGTACTAA CCCACGATTC TGAGCCCTGA GTATGCCTGG ACATTGATGC    1020

TAACATGACC ATGCTTGGGA TGTCTCTAGC TGGTCTGGGG ATAGCTGGAG CACTTACTCA    1080

GGTGGCTGGT GAAATGACAC CTCAGAAGGA ATGAGTGCTA TAGAGAGGAG AGAGGAGTGT    1140

ACTGCCCAGG TCTTTGACAG ATGTAATTCT CATTCAATTA AAGTTTCAGT GTTTTGGTTA    1200

AGTGG                                                              1205
```

(2) INFORMATION FOR SEQ ID NO:8:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 455 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..455
        (D) OTHER INFORMATION: /note= "cDNA clone 20sa7 for a homolog
                            of rat gene BEM-1"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:8:

```
GAAATCAGAA GTTTAATATG ACACAATTAA ATATATTTGT ATATCTCACA CCGGAGNTTC       60

TCTTCAAACA TAAGGAGTTA GAAATTACAA GTAGGCATAT GCTTCCTATA TTCAGATAAA      120

TTCATTTCGA TTAATTAAAT TCCAGATAGA GAGAAGTAAT TTTCGGAAAA GAAATGATAG      180

CTATATTAAA GCAGATATTC ATTACAATAC CATGTAGAGA CATAAGCAAT ATTTTGGCAT      240

CATTCTGTCC GCTCAGTAGG CCGTGTTCCC TCTGGTAGGG CCTTTGGAGA GTACCATCTA      300

TCTAAGATGG AGGAATGCTG TGGGAAGGGC GGGATGGAGG TGCGTTTTCT ACGCTGAACC      360

CCACACAGGA AATCTGCAGC CCACACAGCT GCCTCTGCGC CGCCTTCCAT GTGATCATCC      420

TGGTCAATGA AGTGAATTGT CCTATTTCNG GGGGT                                455
```

(2) INFORMATION FOR SEQ ID NO:9:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 10365 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..10365
        (D) OTHER INFORMATION: /note= "genomic sequence encoding zinc
                       finger amplified in breast cancer
                       (ZABC-1) gene"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:9:

```
CCATCATATT TCTTATTTTT TTGGGCGGAG AGGGGAGACT TGCTCTGTTG CCCAGGCTGG       60

ACCAGTGGTG CGATCTTGGC TCACTGCAAC CTCCACCTCC TGGGTTCAAG TGATTCCCAA      120

ATAGCTGGGA TTACAGGTGT GTATTACCAT GCCCAGCTAA TTTTTGTATT TTTAGCAGAT      180

AAGGGGTTTC ACCATGTTGG CCAGGCTGGT CTCCAACTCC TGGCCTCATG TGATCCACCC      240

ACTTCGGCTT CCCAAAGCAT TGGGAGTATA GGTGTGAGCC ACTATACCCG TCCTCACATC      300

ATATTTCTAA TCCCGAGACT GTAGAGCTGG TGTCTCTTTT TCTAAAGGAT GTCAGTAGAG      360

AAGTGGAGTT CCCCAAAATT ACAGTTTCAC GTATTAGTCA AGTTTCTAAA ATACAGTAAT      420

AATGTTGAGA GCTGACATAG GGACTAACTT GGTTTTTTTT TTTTTTTTTT TTTTTCAAAT      480

TCTCACTGAA CTTTGATTTT GCTAAATAAG GACATTAAAA AAAAAACCAA AAAACTCCAC      540

TATTGCCTAT TGCCACTATT TGATTTTTTA AAAAATAAGC GTATTTTAGC ATCTAAAAGT      600

AGGAAGGACC TCAAATAAAT GAGTCTTTGT TCTTGGCCAG GGAAAACAGC GTTGTCAGAA      660

TTTGATAACT GTTTTTCTAG GGTATGTGCT GTTATTCAGT TAAAACCTTG CCTGGGACGC      720

TAGCATTCAG TAAATACTTG TTGAATAAGC AAATGAAACT TAAGCTTCTA TGTATAGAAA      780

CCTAAGTCAC TTCACATTCT GATTAGCAGA GTAATTGAAT ATTCTTTTCA ATGTGTAGCT      840
```

```
CTATCCCCAG  AACCACAGAA  TATTGGAACT  GTAAAGGCCA  TCCTATAGTT  TAACCAACTG    900

CGTTAAATAG  ATAATAGAAA  GATGTGGTAT  GTGGCAGTGA  CAACTTGAAG  GTTGTGACTA    960

GAACTCGGGT  CTCTGGAGTG  TTCTATTATA  TCACACCAAG  CTGGTCACCA  GCCCATGTGT   1020

TGATCCTCCA  TTGTGATAGC  AACAAAGAAA  AGACTTCAGG  ACATTCTTTC  CTTTACCCTA   1080

ATCCTTGATC  TGCAGTCTTA  TTTAGAAAAG  CTTAATGTTA  AAGATCTAGT  TTATTCAAAA   1140

CTAAAGATAA  CAAGGAGTAT  GAGAATTTCT  ATTTCGGAGT  GTAAAGGAGG  AGATGTTTCC   1200

TTGGCTTCTC  TGAGCCTGCA  GGCCTTCCTT  GCTCTTTAAG  GAAGTAGAGA  GAGGGAGGAA   1260

AGTAAAGTAT  GCTTTTGTTT  TTTAAGGTTA  CTTTGCTGGG  AGTAGTTTGC  ATGCCTTTTG   1320

GTTTTCTTGG  GTGGAATTAA  CTGACTTAAG  TTTTAAGTAG  TTGGGACTAT  TTAAAAACAA   1380

TGCCTATCCA  ATGTTTGCCA  TAAAGGCAGA  GGGTATTGGC  TTTAGAAGTT  AATTCTTCTC   1440

CAGGAGTGAA  AATTAGCTTC  TAAACCAGAA  GCAGCAGAGC  TAAATAAAGT  AATTTTCCAC   1500

CTGGCCAGTG  CATGATGTGA  AAGGTAGATT  AAAAAAATGA  GAGGGCCCAT  TTTCTGATGA   1560

AAGACTAAGC  CATGTTGAAA  CAGCCCTGTT  GAGGATTTTA  TTTTAAATCT  ATACATTCAC   1620

AAAGGAGCTT  TGTGTATGTC  TTTCCCTATT  TGTTGTTTGG  ACTAGGAAGC  CCCACCCAGT   1680

GCTTGTTGAA  GGCAGAAAGT  CGTTGAAAGC  AAGCTGGGAT  TTGAACAGTG  GATTGAGGTT   1740

TCGAATATCC  AGTGAACCAA  AATATATCAG  GGTTCCCCTG  GCCAAGATGA  GTGACCATTC   1800

TGAGGTGTTA  AGTATTTCTT  GAATGGGGAT  TTTAGGAAAA  GTTTCTGTAT  TTCTGTGCTC   1860

ATTTTGTTGA  CCTCTGTATG  TGCAAAATCT  CTAAGGGGGT  GTTTGGGCAC  TTAGATTTCT   1920

TGGATGCAGA  TTTGTTTGTA  TATGAAACAA  ATTTTAAATT  GTTTTGTATA  CACTGGATTT   1980

AAAATAGTTT  ACTAAAGTGT  TTTAATTTTT  TCATCTTAAT  TTTCACAGTT  CTTATAGTCT   2040

TTAGATTTAG  GGAGGCTGTT  GATGGCATCC  ACATGTGCAT  TTTAGTGGCA  TTTAAAATGT   2100

ATTCAGCTGA  ATTTAACAAT  TTCTGACCTA  AAACTTGACA  TTTTAGATTT  AAGTCGGTAA   2160

AGCACTGATT  TAAACTGGAT  TTTAACTGGA  TGAAATTCTG  ATTTAATAAG  TGTACTGACT   2220

GGATAAAATG  CCAATGATTT  AATTAACAAG  CACGTTTAAC  AGGATGCCCT  ATATATTAGT   2280

TAAAAGTGAA  GCAATTGAAT  TAGGTACCTT  CTCTGCTGCG  TGGAAAAGAC  CGTATGACTC   2340

ACCCACACCA  GCCTTCTCTT  CGCTCTGAGT  GTAGCTAACC  GTTTCTGTTT  TTTTTCCTCT   2400

AGGGTTTGGA  AATCCCTTGT  CTCCAGGTTG  CTGGGATTGA  CTTCTTGCTC  AATTGAAACA   2460

CTCATTCAAT  GGAGACAAAG  AGAACTAATG  CTTTGTGCTG  ATTCATATTT  GAATCGAGGC   2520

ATTGGGAACC  CTGTATGCCT  TGTTTGTGGA  AAGAACCAGT  GACACCATCA  CTGAGCTTCC   2580

TAAAAGTTCG  AAGAAGTTAG  AGGACTATAC  ACTTTCTTTT  GAACTTTTAT  AATAAATATT   2640

TGCTCTGGTT  TTTGGAACCC  AGGGCTGTTA  GAGGGGTGAG  TGACAAGTCT  TACAAGTGGC   2700

CTTATTCCAA  CTCCAGAAAT  TGCCCAACGG  AACTTTGAGA  TTATATGCAA  TCGAAAGTGA   2760

CAGGAAACAT  GCCAACTCAA  TCCCTCTTAA  TGTACATGGA  TGGGCCAGAA  GTGATTGGCA   2820

GCTCTCTTGG  CAGTCCGATG  GAGATGGAGG  ATGCCTTGTC  AATGAAAGGG  ACCGCTGTTG   2880
```

TTCCATTCCG AGCTACACAA GAAAAAAATG TCATCCAAAT CGAGGGGTAT ATGCCCTTGG   2940

ATTGCATGTT CTGCAGCCAG ACCTTCACAC ATTCAGAAGA CCTTAATAAA CATGTCTTAA   3000

TGCAACACCG GCCTACCCTC TGTGAACCAG CAGTTCTTCG GGTTGAAGCA GAGTATCTCA   3060

GTCCGCTTGA TAAAAGTCAA GTGCGAACAG AACCTCCCAA GGAAAAGAAT TGCAAGGAAA   3120

ATGAATTTAG CTGTGAGGTA TGTGGGCAGA CATTTAGAGT CGCTTTTGAT GTTGAGATCC   3180

ACATGAGAAC ACACAAAGAT TCTTTCACTT ACGGGTGTAA CATGTGCGGA AGAAGATTCA   3240

AGGAGCCTTG GTTTCTTAAA AATCACATGC GGACACATAA TGGCAAATCG GGGGCCAGAA   3300

GCAAACTGCA GCAAGGCTTG GAGAGTAGTC CAGCAACGAT CAACGAGGTC GTCCAGGTGC   3360

ACGCGGCCGA GAGCATCTCC TCTCCTTACA AAATCTGCAT GGTTTGTGGC TTCCTATTTC   3420

CAAATAAAGA AAGTCTAATT GAGCACCGCA AGGTGCACAC CAAAAAAACT GCTTTCGGTA   3480

CCAGCAGCGC GCAGACAGAC TCTCCACAAG GAGGAATGCC GTCCTCGAGG GAGGACTTCC   3540

TGCAGTTGTT CAACTTGAGA CCAAAATCTC ACCCTGAAAC GGGGAAGAAG CCTGTCAGAT   3600

GCATCCCTCA GCTCGATCCG TTCACCACCT TCCAGGCTTG GCAGCTGGCT ACCAAAGGAA   3660

AAGTTGCCAT TTGCCAAGAA GTGAAGGAAT CGGGGCAAGA AGGGAGCACC GACAACGACG   3720

ATTCGAGTTC CGAGAAGGAG CTTGGAGAAA CAAATAAGGG CAGTTGTGCA GGCCTCTCGC   3780

AAGAGAAAGA GAAGTGCAAA CACTCCCACG GCGAAGCGCC CTCCGTGGAC GCGGATCCCA   3840

AGTTACCCAG TAGCAAGGAG AAGCCCACTC ACTGCTCCGA GTGCGGCAAA GCTTTCAGAA   3900

CCTACCACCA GCTGGTCTTG CACTCCAGGG TCCACAAGAA GGACCGGAGG GCCGGCGCGG   3960

AGTCGCCCAC CATGTCTGTG GACGGGAGGC AGCCGGGGAC GTGTTCTCCT GACCTCGCCG   4020

CCCCTCTGGA TGAAAATGGA GCCGTGGATC GAGGGGAAGG TGGTTCTGAA GACGGATCTG   4080

AGGATGGGCT TCCCGAAGGA ATCCATCTGG GTAAGCTGCC CTGTCTCCGT CCCGTGCTGT   4140

TCCGCCTGTG TCTGTCTGTC TCCCCGTCTC CCCCTCTCTA TTCCCATCTC CAGACAACGC   4200

TGGCCAGGAA TGGGGTTTGG AGAGCCAGAG TCAAGTCCAG GCTCTTTTTG GTATCACTCT   4260

GTGTAAGTCA TTTAACCTCT CAGGGCCTTA ATTTTCTCAT TTCTGTAATA ACAGGGTTGA   4320

GTTAAGAGGT CTCCTTGTTC TGAAAATATA TATATATTTT TTAAACGTGT ATCGTTTTGC   4380

TCACAAAACA CACTTTAAAA AAAAAATAAC TTGTGCATCC AGCCCAAATG CACTGCTTCT   4440

TAACTGGGGC GATTTTGTTC CCAATCAGTA TCTGGCAATG TCTGGAGGCA TTTTGGTTGT   4500

CATACTGTGT GTGTGGGTGT GCCTGCTGGC ATCCAGTGGG CAGAGGCCAG GGACACTGCT   4560

CAGCATGGTA CAGTGCACAG GACAGCCCCA TCATCAAAGA ATTATCTGGT CCCAAATGTC   4620

AATAGTTTGA GCATTGAGAG ACCCTAGCCT TCACTTAAGT TTTTCTGGCG TTCCTGATCT   4680

TTTTCTGTAG TGAATTTCTA GTGGCCATAA AAGGTACTGG GAGTGATCAA CTAGAGCCAG   4740

GAATATTATT TGGGCAGCCG TTTGGTGCTG TCCAAAACCT TGTCCTTTCT GTCTGGCAAG   4800

CTAGTATCCA TTTATAGGTA CCTCAGGAAC CCAAATGATT TGTCATAAAA TACAAGGAAT   4860

GTGAGCACAC TGAAGACATT TTTAAGAAGG CTCATTTGCT CAGCAGAATT TTCAGTGTAC   4920

```
TAGTGGCATT TATAGAAAGA GAAGGTGATC ACTGAAGGCA TGCTCACATA ATATTCCTGA    4980

GCCCTGGTGG GCGTTATCTA GGGCAAAGGA TTCCACCTGT GTTTGGAGTT GCGCCCATCC    5040

TCACTGTAGC CAGAGCTTCT CCTATCAGAG TTTAGTATTT TGTTTGAATA GAGGATCTTG    5100

CTGCTTAAAA CAGTTGAAAA GACCCTGATG GGCAGGCCGT AATTGACAAG CGAATGATGG    5160

GAACATGAAT CGGTCTTAGG GAAGCATCTG TCAAAGTGGT CCTTGGTTAA AACAAGTGCC    5220

TCCTCCTCTC AGTGTCACTT GATTGTGTGC TTGAATTCTT CGGAAAACTG GGTGTATGAG    5280

ACCCACGATG AATTTGCCCA CACGATTGAT TGGACTCTTC CTTCACCTGC TCTTCAGCCA    5340

GTGCCAGTTC CTTTTCTGAT CATGTGATTG ACGTGAGAAC TGTAGTCTGT ATATCAAATC    5400

TTTAGAATGT TTTTGAGTTT CCTGGGACAC AGGAAACCCA GCACTTAGCA TACTACAAAT    5460

CTAATGTCTT AATGGCATCA TAAAAGAGG CTTTAAACAC AGACTCCAGT TAGCTAAGTG    5520

GTTTCTGCTA GTGCCGGTAC TGTTGCAGGG GCCCTGTGAG ATGCCCCAGT TCCCTGAAAG    5580

AAATGAAAAG GCCAGTTACC GGTAGGTGGT GTGGAAAACA TGGGCTAGAT CATCAGGCAG    5640

GACAGAATGC CTGGCTGTGG GTGGGAGCAC CCCAGCTTGG CGTTGAGTTC TGGTTCTACC    5700

ACTGCGTTGT TTTGTGACCA ATTATGAGTT GCTTAACCTT TCTTTGCTAC TATTTCCCTG    5760

TTTGCAAAAT GGTTCATTGA CCCCTGTCTT CCACCTCCCA AGGACAATTT CAACAGCCTA    5820

TTTGTAAAAA GATCACAGTC CTTTAAAAAA TATAACTGTA AAGTCAGAGG TGATGCTTGA    5880

AAGAGCAGGA ACCAGGTAGA TGTGGAAATG TCATGTCCTT TGTTCTAAAG AAAAGGCATT    5940

TCATAGCTTT TTGGATATGA CGCAACATAC CATAAATCCT GACACATAGT TGGGAGTCGG    6000

AAATTGCAAC AACGCCCAGT TATAAACCCA GCTAGTTTGG GTATGATTGT AAGAAAAAAA    6060

AGCTGGCCAT TCTGTATTTG GGGAATTGAT TTTCCTAAAC TTATATTATC TTAGTAGTCT    6120

AGATTTATCA TATTGTACTA TCATCCTGGC TTTTTTAAGA CTTAAGAAGA TCAAGTAAAT    6180

TTTTTTTTCT TTCTTTAGAC ACTATATAGA TCATCAAGGG TGTCTGTCTT ACAGGTGGAT    6240

AGTGATATGA TCTACAGTGA GGGGACATTT ATTTAAAACT TAAACATTCA TGTGTTTTGG    6300

GGGTGGTATT TTAACGGCAG CACCTCTGAT TGTCTTTTGG AGGGCTGGTG TGTGTTTGAA    6360

GTTCTGTCCT CCTTCCAGTG GACTCTAACT TCTCCTGATG CACGTGAGAC ACATTGTCCT    6420

ATTGTCCTGC AGAAACTAAA GCCAAACACT GTCATCTGGG GACAGGTTTT CATTTGTCAG    6480

ATCTCTTTCG CCCACATGAG TGTTTGTGGA CAATACAGCC TGCTTTCCAA AACTTTGCTA    6540

AATTTTGACA GACTTTCCTA GGTGCTTGCC CAATGCCAGA CTTTCTTTTC TGTTGAAGAT    6600

TAAGTTGTGC TTGCTGCCCT CTAGTGGTCA GTTGTTTAAT CCTAACCTTA AACGGCTTAT    6660

TTTTCCCCTG GTGGTTGGGA AGTTGACGGT TTGTAATTGG CTCATTTTTC TAAATTATTC    6720

TGAAGAAGAT AATTTTTCCC GCCAGTATGT ATGTCCACCT TCAGTTTGCC AGATCCTGCC    6780

TGCTCAGAGA CACTGAGAAC CGGAAGCTGC CCGGGCAATT CAGTCTATGA AATGATCTTT    6840

CTTGTGATTA AGGCAAACGA AGAACTGAAT GTTAATAGT GTACTCTGCT GTACCCAGAA    6900

AAAAACAAAA CAAAATCATG TTATAACACT CTAAAACTTC AAACAACCTC CAACAGCATT    6960
```

```
TGGTGTGTGT CTAGCCGTTT TGTTCTAACC CGATGTTATA TAAAAGAATT TTTTCATGCT    7020

TTCCAAAAAT GTTTATGTCA AGAATATTTA AGTCAGCATG CCTTATTCAG GTACTTCAGC    7080

TACCTTCTTA TATAAATATT TTTGTTTTTC CTTTAAGATA AAAATGATGA TGGAGGAAAA    7140

ATAAAACATC TTACATCTTC AAGAGAGTGT AGTTATTGTG GAAAGTTTTT CCGTTCAAAT    7200

TATTACCTCA ATATTCATCT CAGAACGCAT ACAGGTAAAG AACTTTTATT TTTTTAACCA    7260

TGCATTAGTT AAATTATGTA GTTATCTAAT TTTTTTGTTG TTGTTGTTCA GATACTCTGC    7320

CAGATCCTTG GACTAGCTTA AGGATAAATA TGTAGCATGT TGATTGCAGT GGTTATTTTT    7380

ATTCTTTTAG TGCCATTGTA ACTTGAGCCA TTGTTCTTAT TTGCAGTTCA TTTCTTTTCT    7440

TTCTTTTTTG TTTTTTGAGA CGGAGTCTTG CTCTGTCACC TCGGCTGGAG TGCAGTGGTG    7500

CAATTTCGGC TCACTGCAGC CTCCACCTCC CTGGTTCAAG CAATACTCCT GCCTCAGCCT    7560

CCCCAGTAGT TGGGATTACA GGTACCTGCC ACCACACCCG GCTAATTTCT GTATTTTTAG    7620

TAGAGATGGG GTTTCACCAT GCTGGCCAGG CTGGTTTCGA ACTCCTGACC TCAAGTGATC    7680

CGCTCACCTT GGCCTCCCAT AGTGTTGGCC TCCCATAGTG CTGGGATTAC AGGCGTGAGC    7740

CACCGCGCCC GGACAAAGTT CATTTGTTTA GTTTATGACT GCTATGTCCT GACTCTTATC    7800

TTATTAAAAG CTACAGTATT TTAAAATGCT GCATCTTATG TCTTTATGAT TGAGAATGAA    7860

ATGAGAATCT ATTTAGTAGT CTTGAGATTG TGAAAGGAGC TATGACATCA TGATGTAGGA    7920

GGCTGCGTAG ATTTGAAATT TCATCTCTTC CACTTACTAT CTGTGCACCC TTGGGCAAGT    7980

TATTTAACCT TTTTGTGCTT TTAGTTTTCT TTGCTGTAAA AGTAGAATAA TACATATTTC    8040

CCTAGGGCTG TTAGGAAGAT TAAATAAGTT AGAAGTGTTG CTGTTAATTT TTCTATTGAA    8100

GATAGGCATT CATAATTTCA AATATTCATT ACAGTAAGGA TGATAAAGAA CTGATGAGAA    8160

ATCCTATGTG ATAGTAGATC GAGAAAGCAA AAGGAGGAAA GAAGCCTGTT TTCTTAATAA    8220

ATAGATATTT GATCTATTTC AGTGCTTTTC ATACACTTCT ATAATAAAGT GCCATTTCTT    8280

GCCTTAGGTG AAAAACCATA CAAATGTGAA TTTTGTGAAT ATGCTGCAGC CCAGAAGACA    8340

TCTCTGAGGT ATCACTTGGA GAGACATCAC AAGGAAAAAC AAACCGATGT TGCTGCTGAA    8400

GTCAAGAACG ATGGTAAAAA TCAGGACACT GAAGATGCAC TATTAACCGC TGACAGTGCG    8460

CAAACCAAAA ATTTGAAAAG ATTTTTTGAT GGTGCCAAAG ATGTTACAGG CAGTCCACCT    8520

GCAAAGCAGC TTAAGGAGAT GCCTTCTGTT TTTCAGAATG TTCTGGGCAG CGCTGTCCTC    8580

TCACCAGCAC ACAAAGATAC TCAGGATTTC CATAAAAATG CAGCTGATGA CAGTGCTGAT    8640

AAAGTGAATA AAAACCCTAC CCCTGCTTAC CTGGACCTGT TAAAAAAGAG ATCAGCAGTT    8700

GAAACTCAGG CAAATAACCT CATCTGTAGA ACCAAGGCGG ATGTTACTCC TCCTCCGGAT    8760

GGCAGTACCA CCCATAACCT TGAAGTTAGC CCCAAAGAGA AGCAAACGGA GACCGCAGCT    8820

GACTGCAGAT ACAGGCCAAG TGTGGATTGT CACGAAAAAC CTTTAAATTT ATCCGTGGGG    8880

GCTCTTCACA ATTGCCCGGC AATTTCTTTG AGTAAAAGTT TGATTCCAAG TATCACCTGT    8940

CCATTTTGTA CCTTCAAGAC ATTTTATCCA GAAGTTTTAA TGATGCACCA GAGACTGGAG    9000
```

```
CATAAATACA ATCCTGACGT TCATAAAAAC TGTCGAAACA AGTCCTTGCT TAGAAGTCGA      9060

CGTACCGGAT GCCCGCCAGC GTTGCTGGGA AAAGATGTGC CTCCCCTCCC TAGTTTCTGT      9120

AAACCCAAGC CCAAGTCTGC TTTCCCGGCG CAGTCCAAAT CCCTGCCATC TGCGAAGGGG      9180

AAGCAGAGCC CTCCTGGGCC AGGCAAGGCC CCTCTGACTT CAGGGATAGA CTCTAGCACT      9240

TTAGCCCCAA GTAACCTGAA GTCCCACAGA CCACAGCAGA ATGTGGGGGT CCAAGGGGCC      9300

GCCACCAGGC AACAGCAATC TGAGATGTTT CCTAAAACCA GTGTTTCCCC TGCACCGGAT      9360

AAGACAAAAA GACCCGAGAC AAAATTGAAA CCTCTTCCAG TAGCTCCTTC TCAGCCCACC      9420

CTCGGCAGCA GTAACATCAA TGGTTCCATC GACTACCCCG CCAAGAACGA CAGCCCGTGG      9480

GCACCTCCGG GAAGAGACTA TTTCTGTAAT CGGAGTGCCA GCAATACTGC AGCAGAATTT      9540

GGTGAGCCCC TTCCAAAAAG ACTGAAGTCC AGCGTGGTTG CCCTTGACGT TGACCAGCCC      9600

GGGGCCAATT ACAGAAGAGG CTATGACCTT CCCAAGTACC ATATGGTCAG AGGCATCACA      9660

TCACTGTTAC CGCAGGACTG TGTGTATCCG TCGCAGGCGC TGCCTCCCAA ACCAAGGTTC      9720

CTGAGCTCCA GCGAGGTCGA TTCTCCAAAT GTGCTGACTG TTCAGAAGCC CTATGGTGGC      9780

TCCGGGCCAC TTTACACTTG TGTGCCTGCT GGTAGTCCAG CATCCAGCTC GACGTTAGAA      9840

GGTATTGCAT GAGGGGCGTC GTGTTTAAAT GGCTGCCTAC AGTGATTAAT AGCTAATCCA      9900

GGCATTCTCA GTGGAGATGG TACCACTCCC AAGGGTGGGG GGTAGGCAGC CAGAAGTTCT      9960

TGGGGGTCAC AGAGAGAAGC ATTCTTAGAT ACGGCAGTGG TTTGTGGTCC TCCAAGGCTT     10020

ACTTAACTCT GTGGGTTTAA CTCTTAACCC TGTGTATTTT ATTCTTTTGA TTTGTTTAGT     10080

CTTACTTTAT TTTTAGAGAA AGGGTCTTGC TCCGTCATCT AGATTGGAGT GCAGCGGTGT     10140

AATCATAGCT TACTGTAGTC TTGAATTCCT GAGTTCAAGA GATCCTTCTG CCTCAGCTTC     10200

CCAGGTAGCT GAGACTATAT GTGCTGCTAC CATGCACAGC TGATTTTTAA ATTTTTTTTG     10260

TAGAGATGGA GTTGCCCAGG CTGGTCTTGA ACTCCTGGCC TGAGGTGATC CTCCTGCGTT     10320

GACCTCCCAA GTATCTTAGA CTACAGATGC ACTCCACCAC GCTTG                     10365
```

(2) INFORMATION FOR SEQ ID NO:10:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 3186 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..3186
        (D) OTHER INFORMATION: /note= "ZABC-1 open reading frame"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:10:

```
ATGCAATCGA AAGTGACAGG AAACATGCCA ACTCAATCCC TCTTAATGTA CATGGATGGG      60

CCAGAAGTGA TTGGCAGCTC TCTTGGCAGT CCGATGGAGA TGGAGGATGC CTTGTCAATG     120

AAAGGGACCG CTGTTGTTCC ATTCCGAGCT ACACAAGAAA AAAATGTCAT CCAAATCGAG     180

GGGTATATGC CCTTGGATTG CATGTTCTGC AGCCAGACCT TCACACATTC AGAAGACCTT     240

AATAAACATG TCTTAATGCA ACACCGGCCT ACCCTCTGTG AACCAGCAGT TCTTCGGGTT     300

GAAGCAGAGT ATCTCAGTCC GCTTGATAAA AGTCAAGTGC GAACAGAACC TCCCAAGGAA     360

AAGAATTGCA AGGAAAATGA ATTTAGCTGT GAGGTATGTG GGCAGACATT TAGAGTCGCT     420

TTTGATGTTG AGATCCACAT GAGAACACAC AAAGATTCTT TCACTTACGG GTGTAACATG     480

TGCGGAAGAA GMTTSRRSSA GCCTTGGTTT CTTAAAAATC ACATGCGGAC ACATAATGGC     540

AAATCGGGGG CCAGAAGCAA ACTGCAGCAA GGCTTGGAGA GTAGTCCAGC AACGATCAAC     600

GAGGTCGTCC AGGTGCACGC GGCCGAGAGC ATCTCCTCTC CTTACAAAAT CTGCATGGTT     660

TGTGGCTTCC TATTTCCAAA TAAAGAAAGT CTAATTGAGC ACCGCAAGGT GCACACCAAA     720

AAAACTGCTT TCGGTACCAG CAGCGCGCAG ACAGACTCTC CACAAGGAGG AATGCCGTCC     780

TCGAGGGAGG ACTTCCTGCA GTTGTTCAAC TTGAGACCAA AATCTCACCC TGAAACGGGG     840

AAGAAGCCTG TCAGATGCAT CCCTCAGCTC GATCCGTTCA CCACCTTCCA GGCTTGGCAG     900

CTGGCTACCA AAGGAAAAGT TGCCATTTGC CAAGAAGTGA AGGAATCGGG GCAAGAAGGG     960

AGCACCGACA ACGACGATTC GAGTTCCGAG AAGGAGCTTG GAGAAACAAA TAAGGGCAGT    1020

TGTGCAGGCC TCTCGCAAGA GAAAGAGAAG TGCAAACACT CCCACGGCGA AGCGCCCTCC    1080

GTGGACGCGG ATCCCAAGTT ACCCAGTAGC AAGGAGAAGC CCACTCACTG CTCCGAGTGC    1140

GGCAAAGCTT TCAGAACCTA CCACCAGCTG GTCTTGCACT CCAGGGTCCA CAAGAAGGAC    1200

CGGAGGGCCG GCGCGGAGTC GCCCACCATG TCTGTGGACG GGAGGCAGCC GGGGACGTGT    1260

TCTCCTGACC TCGCCGCCCC TCTGGATGAA AATGGAGCCG TGGATCGAGG GGAAGGTGGT    1320

TCTGAAGACG GATCTGAGGA TGGGCTTCCC GAAGGAATCC ATCTGGATAA AAATGATGAT    1380

GGAGGAAAAA TAAAACATCT TACATCTTCA AGAGAGTGTA GTTATTGTGG AAAGTTTTTC    1440

CGTTCAAATT ATTACCTCAA TATTCATCTC AGAACGCATA CAGGTGAAAA ACCATACAAA    1500

TGTGAATTTT GTGAATATGC TGCAGCCCAG AAGACATCTC TGAGGTATCA CTTGGAGAGA    1560

CATCACAAGG AAAAACAAAC CGATGTTGCT GCTGAAGTCA AGAACGATGG TAAAAATCAG    1620

GACACTGAAG ATGCACTATT AACCGCTGAC AGTGCGCAAA CCAAAAATTT GAAAAGATTT    1680

TTTGATGGTG CCAAAGATGT TACAGGCAGT CCACCTGCAA AGCAGCTTAA GGAGATGCCT    1740

TCTGTTTTTC AGAATGTTCT GGGCAGCGCT GTCCTCTCAC CAGCACACAA AGATACTCAG    1800

GATTTCCATA AAAATGCAGC TGATGACAGT GCTGATAAAG TGAATAAAAA CCCTACCCCT    1860

GCTTACCTGG ACCTGTTAAA AAAGAGATCA GCAGTTGAAA CTCAGGCAAA TAACCTCATC    1920

TGTAGAACCA AGGCGGATGT TACTCCTCCT CCGGATGGCA GTACCACCCA TAACCTTGAA    1980
```

```
GTTAGCCCCA AAGAGAAGCA AACGGAGACC GCAGCTGACT GCAGATACAG GCCAAGTGTG    2040

GATTGTCACG AAAAACCTTT AAATTTATCC GTGGGGGCTC TTCACAATTG CCCGGCAATT    2100

TCTTTGAGTA AAAGTTTGAT TCCAAGTATC ACCTGTCCAT TTTGTACCTT CAAGACATTT    2160

TATCCAGAAG TTTTAATGAT GCACCAGAGA CTGGAGCATA AATACAATCC TGACGTTCAT    2220

AAAAACTGTC GAAACAAGTC CTTGCTTAGA AGTCGACGTA CCGGATGCCC GCCAGCGTTG    2280

CTGGGAAAAG ATGTGCCTCC CCTCTCTAGT TTCTGTAAAC CCAAGCCCAA GTCTGCTTTC    2340

CCGGCGCAGT CCAAATCCCT GCCATCTGCG AAGGGGAAGC AGAGCCCTCC TGGGCCAGGC    2400

AAGGCCCCTC TGACTTCAGG GATAGACTCT AGCACTTTAG CCCCAAGTAA CCTGAAGTCC    2460

CACAGACCAC AGCAGAATGT GGGGGTCCAA GGGGCCGCCA CCAGGCAACA GCAATCTGAG    2520

ATGTTTCCTA AAACCAGTGT TTCCCCTGCA .CCGGATAAGA CAAAAAGACC CGAGACAAAA    2580

TTGAAACCTC TTCCAGTAGC TCCTTCTCAG CCCACCCTCG GCAGCAGTAA CATCAATGGT    2640

TCCATCGACT ACCCCGCCAA GAACGACAGC CCGTGGGCAC CTCCGGGAAG AGACTATTTC    2700

TGTAATCGGA GTGCCAGCAA TACTGCAGCA GAATTTGGTG AGCCCCTTCC AAAAAGACTG    2760

AAGTCCAGCG TGGTTGCCCT TGACGTTGAC CAGCCCGGGG CCAATTACAG AAGAGGCTAT    2820

GACCTTCCCA AGTACCATAT GGTCAGAGGC ATCACATCAC TGTTACCGCA GGACTGTGTG    2880

TATCCGTCGC AGGCGCTGCC TCCCAAACCA AGGTTCCTGA GCTCCAGCGA GGTCGATTCT    2940

CCAAATGTGC TGACTGTTCA GAAGCCCTAT GGTGGCTCCG GGCCACTTTA CACTTGTGTG    3000

CCTGCTGGTA GTCCAGCATC CAGCTCGACG TTAGAAGGTC TTGGTGGATG TCAGTGCTTA    3060

CTCCCCATGA AATTAAATTT TACTTCATCC TTTGAGAAGC GAATGGTGAA AGCTACTGAA    3120

ATAAGCTGTG ATTGTACTGT ACATAAAACA TATGAGGAAT CTGCAAGGAA CACTACAGTT    3180

GTGTAA                                                              3186
```

(2) INFORMATION FOR SEQ ID NO:11:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 1061 amino acids
        (B) TYPE: amino acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: protein

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..1061
        (D) OTHER INFORMATION: /note= "ZABC-1 protein"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:11:

```
Met Gln Ser Lys Val Thr Gly Asn Met Pro Thr Gln Ser Leu Leu Met
1               5                   10                  15

Tyr Met Asp Gly Pro Glu Val Ile Gly Ser Ser Leu Gly Ser Pro Met
                20                  25                  30
```

```
Glu Met Glu Asp Ala Leu Ser Met Lys Gly Thr Ala Val Val Pro Phe
        35              40                  45

Arg Ala Thr Gln Glu Lys Asn Val Ile Gln Ile Glu Gly Tyr Met Pro
    50              55                  60

Leu Asp Cys Met Phe Cys Ser Gln Thr Phe Thr His Ser Glu Asp Leu
65              70                  75                      80

Asn Lys His Val Leu Met Gln His Arg Pro Thr Leu Cys Glu Pro Ala
                85              90                  95

Val Leu Arg Val Glu Ala Glu Tyr Leu Ser Pro Leu Asp Lys Ser Gln
            100             105             110

Val Arg Thr Glu Pro Pro Lys Glu Lys Asn Cys Lys Glu Asn Glu Phe
        115             120                 125

Ser Cys Glu Val Cys Gly Gln Thr Phe Arg Val Ala Phe Asp Val Glu
    130             135                 140

Ile His Met Arg Thr His Lys Asp Ser Phe Thr Tyr Gly Cys Asn Met
145                 150                 155                 160

Cys Gly Arg Xaa Xaa Xaa Xaa Pro Trp Phe Leu Lys Asn His Met Arg
            165             170                 175

Thr His Asn Gly Lys Ser Gly Ala Arg Ser Lys Leu Gln Gln Gly Leu
            180             185                 190

Glu Ser Ser Pro Ala Thr Ile Asn Glu Val Val Gln Val His Ala Ala
    195             200                 205

Glu Ser Ile Ser Ser Pro Tyr Lys Ile Cys Met Val Cys Gly Phe Leu
    210             215                 220

Phe Pro Asn Lys Glu Ser Leu Ile Glu His Arg Lys Val His Thr Lys
225             230                 235                 240

Lys Thr Ala Phe Gly Thr Ser Ser Ala Gln Thr Asp Ser Pro Gln Gly
            245             250                 255

Gly Met Pro Ser Ser Arg Glu Asp Phe Leu Gln Leu Phe Asn Leu Arg
            260             265                 270

Pro Lys Ser His Pro Glu Thr Gly Lys Lys Pro Val Arg Cys Ile Pro
    275                 280                 285

Gln Leu Asp Pro Phe Thr Thr Phe Gln Ala Trp Gln Leu Ala Thr Lys
    290             295                 300

Gly Lys Val Ala Ile Cys Gln Glu Val Lys Glu Ser Gly Gln Glu Gly
305             310                 315                 320

Ser Thr Asp Asn Asp Asp Ser Ser Ser Glu Lys Glu Leu Gly Glu Thr
            325                 330                 335

Asn Lys Gly Ser Cys Ala Gly Leu Ser Gln Glu Lys Glu Lys Cys Lys
            340             345                 350

His Ser His Gly Glu Ala Pro Ser Val Asp Ala Asp Pro Lys Leu Pro
            355             360                 365

Ser Ser Lys Glu Lys Pro Thr His Cys Ser Glu Cys Gly Lys Ala Phe
    370             375                 380
```

Arg Thr Tyr His Gln Leu Val Leu His Ser Arg Val His Lys Lys Asp
385   390   395   400

Arg Arg Ala Gly Ala Glu Ser Pro Thr Met Ser Val Asp Gly Arg Gln
              405   410   415

Pro Gly Thr Cys Ser Pro Asp Leu Ala Ala Pro Leu Asp Glu Asn Gly
              420   425   430

Ala Val Asp Arg Gly Glu Gly Gly Ser Glu Asp Gly Ser Glu Asp Gly
              435   440   445

Leu Pro Glu Gly Ile His Leu Asp Lys Asn Asp Asp Gly Gly Lys Ile
    450   455   460

Lys His Leu Thr Ser Ser Arg Glu Cys Ser Tyr Cys Gly Lys Phe Phe
465   470   475   480

Arg Ser Asn Tyr Tyr Leu Asn Ile His Leu Arg Thr His Thr Gly Glu
              485   490   495

Lys Pro Tyr Lys Cys Glu Phe Cys Glu Tyr Ala Ala Ala Gln Lys Thr
              500   505   510

Ser Leu Arg Tyr His Leu Glu Arg His His Lys Glu Lys Gln Thr Asp
              515   520   525

Val Ala Ala Glu Val Lys Asn Asp Gly Lys Asn Gln Asp Thr Glu Asp
    530   535   540

Ala Leu Leu Thr Ala Asp Ser Ala Gln Thr Lys Asn Leu Lys Arg Phe
545   550   555   560

Phe Asp Gly Ala Lys Asp Val Thr Gly Ser Pro Pro Ala Lys Gln Leu
              565   570   575

Lys Glu Met Pro Ser Val Phe Gln Asn Val Leu Gly Ser Ala Val Leu
              580   585   590

Ser Pro Ala His Lys Asp Thr Gln Asp Phe His Lys Asn Ala Ala Asp
    595   600   605

Asp Ser Ala Asp Lys Val Asn Lys Asn Pro Thr Pro Ala Tyr Leu Asp
    610   615   620

Leu Leu Lys Lys Arg Ser Ala Val Glu Thr Gln Ala Asn Asn Leu Ile
625   630   635   640

Cys Arg Thr Lys Ala Asp Val Thr Pro Pro Asp Gly Ser Thr Thr
              645   650   655

His Asn Leu Glu Val Ser Pro Lys Glu Lys Gln Thr Glu Thr Ala Ala
              660   665   670

Asp Cys Arg Tyr Arg Pro Ser Val Asp Cys His Glu Lys Pro Leu Asn
    675   680   685

Leu Ser Val Gly Ala Leu His Asn Cys Pro Ala Ile Ser Leu Ser Lys
    690   695   700

Ser Leu Ile Pro Ser Ile Thr Cys Pro Phe Cys Thr Phe Lys Thr Phe
705   710   715   720

Tyr Pro Glu Val Leu Met Met His Gln Arg Leu Glu His Lys Tyr Asn
              725   730   735

```
Pro Asp Val His Lys Asn Cys Arg Asn Lys Ser Leu Leu Arg Ser Arg
            740             745             750

Arg Thr Gly Cys Pro Pro Ala Leu Leu Gly Lys Asp Val Pro Pro Leu
            755             760             765

Ser Ser Phe Cys Lys Pro Lys Pro Lys Ser Ala Phe Pro Ala Gln Ser
    770             775             780

Lys Ser Leu Pro Ser Ala Lys Gly Lys Gln Ser Pro Pro Gly Pro Gly
785             790             795             800

Lys Ala Pro Leu Thr Ser Gly Ile Asp Ser Ser Thr Leu Ala Pro Ser
            805             810             815

Asn Leu Lys Ser His Arg Pro Gln Gln Asn Val Gly Val Gln Gly Ala
            820             825             830

Ala Thr Arg Gln Gln Gln Ser Glu Met Phe Pro Lys Thr Ser Val Ser
            835             840             845

Pro Ala Pro Asp Lys Thr Lys Arg Pro Glu Thr Lys Leu Lys Pro Leu
    850             855             860

Pro Val Ala Pro Ser Gln Pro Thr Leu Gly Ser Ser Asn Ile Asn Gly
865             870             875             880

Ser Ile Asp Tyr Pro Ala Lys Asn Asp Ser Pro Trp Ala Pro Pro Gly
            885             890             895

Arg Asp Tyr Phe Cys Asn Arg Ser Ala Ser Asn Thr Ala Ala Glu Phe
            900             905             910

Gly Glu Pro Leu Pro Lys Arg Leu Lys Ser Ser Val Val Ala Leu Asp
            915             920             925

Val Asp Gln Pro Gly Ala Asn Tyr Arg Arg Gly Tyr Asp Leu Pro Lys
    930             935             940

Tyr His Met Val Arg Gly Ile Thr Ser Leu Leu Pro Gln Asp Cys Val
945             950             955             960

Tyr Pro Ser Gln Ala Leu Pro Pro Lys Pro Arg Phe Leu Ser Ser Ser
            965             970             975

Glu Val Asp Ser Pro Asn Val Leu Thr Val Gln Lys Pro Tyr Gly Gly
            980             985             990

Ser Gly Pro Leu Tyr Thr Cys Val Pro Ala Gly Ser Pro Ala Ser Ser
            995             1000            1005

Ser Thr Leu Glu Gly Leu Gly Gly Cys Gln Cys Leu Leu Pro Met Lys
    1010            1015            1020

Leu Asn Phe Thr Ser Ser Phe Glu Lys Arg Met Val Lys Ala Thr Glu
1025            1030            1035            1040

Ile Ser Cys Asp Cys Thr Val His Lys Thr Tyr Glu Glu Ser Ala Arg
            1045            1050            1055

Asn Thr Thr Val Val
            1060
```

(2) INFORMATION FOR SEQ ID NO:12:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 3066 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: cDNA

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..3066
        (D) OTHER INFORMATION: /note= "1b1"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:12:

```
GGAAACAGCT ATGACCATGA TTACGCCAAG CTCGAAATTA ACCCTCACTA AAGGGAACAA      60

AAGCTGGAGC TCCACCGCGG TGGCGGCCGC TCTAGAACTA GTGGATCCCC CGGGCTGCAG     120

GAATTCGGCA CGAGGCTCCA CCGACAGCCA GGCACTGGGC AGCACGCACT GGAGACCCAG     180

GACCCTGTGC AGGAGCAGCT CCGGGTGACA CGAGGGGACT GAAGATACTC CCACAGGGGC     240

TCAGCAGGAG CAATGGGTAA CCAAATGAGT GTTCCCCAAA GAGTTGAAGA CCAAGAGAAT     300

GAACCAGAAG CAGAGACTTA CCAGGACAAC GCGTCTGCTC TGAACGGGGT TCCAGTGGTG     360

GTGTCGACCC ACACAGTTCA GCACTTAGAG GAAGTCGACT TGGGAATAAG TGTCAAGACG     420

GATAATGTGG CCACTTCTTC CCCCGAGACA ACGGAGATAA GTGCTGTTGC GGATGCCAAC     480

GGAAAGAATC TTGGGAAAGA GGCCAAACCC GAGGCACCAG CTGCTAAATC TCGTTTTTTC     540

TTGATGCTCT CTCGGCCTGT ACCAGGACGT ACCGGAGACC AAGCCGCAGA TTCATCCCTT     600

GGATCAGTGA AGCTTGATGT CAGCTCCAAT AAAGCTCCAG CGAACAAAGA CCCAAGTGAG     660

AGCTGGACAC TTCCGGTGGC AGCTGGACCG GGGCAGGACA CAGATAAAAC CCCAGGGCAC     720

GCCCCGGCCC AAGACAAGGT CCTCTCTGCC GCCAGGGATC CCACGCTTCT CCCACCTGAG     780

ACAGGGGGAG CAGGAGGAGA AGCTCCCTCC AAGCCCAAGG ACTCCAGCTT TTTTGACAAA     840

TTCTTCAAGC TGGACAAGGG ACAGGAAAAG GTGCCAGGTG ACAGCCAACA GGAAGCCAAG     900

AGGGCAGAGC ATCAAGACAA GGTGGATGAG GTTCCTGGCT TATCAGGGCA GTCCGATGAT     960

GTCCCTGCAG GGAAGGACAT AGTTGACGGC AAGGAAAAAG AAGGACAAGA ACTTGGAACT    1020

GCGGATTGCT CTGTCCCTGG GGACCCAGAA GGACTGGAGA CTGCAAAGGA CGATTCCCAG    1080

GCAGCAGCTA TAGCAGAGAA TAATAATTCC ATCATGAGTT TCTTTAAAAC TCTGGTTTCA    1140

CCTAACAAAG CTGAAACAAA AAAGGACCCA GAAGACACGG GTGCTGAAAA GTCACCCACC    1200

ACTTCAGCTG ACCTTAAGTC AGACAAAGCC AACTTTACAT CCCAGGAGAC CCAAGGGGCT    1260

GGCAAGAATT CCAAAGGATG CAACCCATCG GGGCACACAC AGTCCGTGAC AACCCCTGAA    1320

CCTGCGAAGG AAGGCACCAA GGAGAAATCA GGACCCACCT CTCTGCCTCT GGGCAAACTG    1380

TTTTGGAAAA AGTCAGTTAA AGAGGACTCA GTCCCCACAG GTGCGGAGGA GAATGTGGTG    1440
```

```
TGTGAGTCAC CAGTAGAGAT TATAAAGTCC AAGGAAGTAG AATCAGCCTT ACAAACAGTG      1500

GACCTCAACG AAGGAGATGC TGCACCTGAA CCCACAGAAG CGAAACTCAA AAGAGAAGAA      1560

AGCAAACCAA GAACCTCTCT GATGGCGTTT CTCAGACAAA TGTCAGTGAA AGGGGATGGA      1620

GGGATCACCC ACTCAGAAGA AATAAATGGG AAAGACTCCA GCTGCCAAAC ATCAGACTCC      1680

ACAGAAAAGA CTATCACACC GCCAGAGCCT GAACCAACAG GAGCACCACA GAAGGGTAAA      1740

GAGGGCTCCT CGAAGGACAA GAAGTCAGCA GCCGAGATGA ACAAGCAGAA GAGCAACAAG      1800

CAGGAAGCCA AGAACCAGC CCAGTGCACA GAGCAGGCCA CGGTGGACAC GAACTCACTG      1860

CAGAATGGGG ACAAGCTCCA AAAGAGACCT GAGAAGCGGC AGCAGTCCCT TGGGGGCTTC      1920

TTTAAAGGCC TGGGACCAAA GCGGATGTTG GATGCTCAAG TGCAAACAGA CCCAGTATCC      1980

ATCGGACCAG TTGGCAAACC CAAGTAAACA AATCAGCACG GTTCCCACCA GGTTCTCCTG      2040

CCACCAAGAT GTGTTCTCCT TACTCCATCT CCTCCCCAAA CACGCTCCAT GTATATATTC      2100

TTCTGATGGC CAGCAAATGA AATTCTGCCT AGAAATTAAG CCCGAGCTGT TGTATATTGA      2160

GGTGTATTAT TTACGTCTCT GGTCCAGTCT TTTCTGGCAA ATAACAGTAA AGATGGTTTA      2220

GCAGGTCACC TAGTTGGGTC AGAAGAGTCG ATGATCACCA AGCAGGAAAG GGAGGGAATA      2280

GAGGAATGTG TTCGGGTTAA GTGATGAAAA TGGCAGTGGT GGCCGGGCGT GGTGGCTCTC      2340

GCCTGTAATC TCAGCACTTT GGGAGGCCGA GGCAGGTGGA TCACCTGAGG TCAGGAGTTC      2400

AAGACTAGCC TGGCCAACAT CATGAAACCC CGTCTCTACT AAAAATACAA AAATTAGCCA      2460

GGCATGGTGG CACACACCTG TAGTCCCAGC TACTCGGGAG CCCAACGCAC GAGAACCGCT      2520

TGTACCCAGG AGGTGGAGGT TGCAGTGAGC CGAAGTTGCA CCATTGCACT CCACCCTGGG      2580

CGACAGAGCA AGATTCTATC AAAAAAAAAA GGCAGTGGCA AGTAAGTTAT AGAAGAGAAA      2640

TGCTGCTAGA AGGAATTAAG CGTTGTAGTA AACGCGTGCT CATCCTCTAA GCTTGAAGAA      2700

GGGAGACGAA AATCCATTTG TTTAAATTCA CATCTCAAGG AGGGAGAACC CGGGCTGTGT      2760

TGGGTGGTTG CCAATTTCCT AGAACGGAAT GTGTGGGGTA TAGAAAAAGG AATGAATAAG      2820

CGTTGTTTTT CAAATAGGGT CCTTGTAAGT TATTGATGAG AGGGAAAAGA TTGACTGGGG      2880

AGGGCTTAAA ATGATTTGGG AAAACAATTG CTTTTGAGGC TCAGTGACAA CGGCAAAGAT      2940

TACAACTTAA AAAAAAAAAA AAAAAACTC GAGACTAGTT CTCTCTCTCT CTCGTGCCGA      3000

ATTCGATATC AAGCTTATCG ATACCGTCGA CCTCGAGGGG GGGCCCGGTA CCCAATTCGC      3060

CCTATA                                                                3066
```

(2) INFORMATION FOR SEQ ID NO:13:

      (i) SEQUENCE CHARACTERISTICS:
          (A) LENGTH: 23 base pairs
          (B) TYPE: nucleic acid
          (C) STRANDEDNESS: single
          (D) TOPOLOGY: linear

      (ii) MOLECULE TYPE: DNA

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:13:

TTGGCATTGG TATCAGGTAG CTG   23

(2) INFORMATION FOR SEQ ID NO:14:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 24 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:14:

TTGGAGCAGA GAGGGGATTG TGTG   24

(2) INFORMATION FOR SEQ ID NO:15:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 24 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:15:

AATCCCCTCA AACCCTGCTG CTAC   24

(2) INFORMATION FOR SEQ ID NO:16:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:16:

TGGAGCCTGA ACTTCTGCAA TC   22

(2) INFORMATION FOR SEQ ID NO:17:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 17 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:17:

CCGGGATACC GACATTG   17

(2) INFORMATION FOR SEQ ID NO:18:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 19 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:18:

TGCACATAAA ACAGCCAGC   19

(2) INFORMATION FOR SEQ ID NO:19:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:19:

TTGGAATCAA TGGAGCAAAA   20

(2) INFORMATION FOR SEQ ID NO:20:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:20:

AGCTTTACCC AATGTGGTCC   20

(2) INFORMATION FOR SEQ ID NO:21:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 21 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:21:

GTGGTGAACA CCAATAAATG G   21

(2) INFORMATION FOR SEQ ID NO:22:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 24 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:22:

AAGCAAATAA AACCAATAAA CTCG   24

(2) INFORMATION FOR SEQ ID NO:23:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 21 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:23:

CAAGATCTGA CCCCGTCAAT C   21

(2) INFORMATION FOR SEQ ID NO:24:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 25 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:24:

GACTTCTTCA GGAAAGAGAT CAGTG   25

(2) INFORMATION FOR SEQ ID NO:25:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 23 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:25:

GCCATGTACC CACCTGAAAA ATC   23

(2) INFORMATION FOR SEQ ID NO:26:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 24 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:26:

TCAGAACACC CGTGCAGAAT TAAG   24

(2) INFORMATION FOR SEQ ID NO:27:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 24 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:27:

CCTAAAACTT GGTGCTTAAA TCTA   24

(2) INFORMATION FOR SEQ ID NO:28:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:28:

GTCTCACAAG GCAGATGTGG   20

(2) INFORMATION FOR SEQ ID NO:29:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:29:

TTTGTGTATG TTGAGCCATC   20

(2) INFORMATION FOR SEQ ID NO:30:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:30:

CTTCCAATCT CATTCTATGA GG   22

(2) INFORMATION FOR SEQ ID NO:31:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:31:

GCTTGTTTAA GTGTCACTAG GG   22

(2) INFORMATION FOR SEQ ID NO:32:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 23 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:32:

CACTCTGGTA AATGACCTTT GTC   23

(2) INFORMATION FOR SEQ ID NO:33:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 21 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:33:

CCTACACCAT TCCAACTTTG G   21

(2) INFORMATION FOR SEQ ID NO:34:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 25 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:34:

GCCAGATGTA TGTTTGCTAC GGAAC   25

(2) INFORMATION FOR SEQ ID NO:35:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:35:

TCTCAAACCT GTCCACTTCT TG   22

(2) INFORMATION FOR SEQ ID NO:36:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 19 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:36:

CTGCTGTGGT GGAGAATGG   19

(2) INFORMATION FOR SEQ ID NO:37:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 24 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:37:

TGTCCTCCTT CTCCCTCATC CTAC   24

(2) INFORMATION FOR SEQ ID NO:38:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 22 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:38:

AATGCCTCCA CTCACAGGAA TG   22

(2) INFORMATION FOR SEQ ID NO:39:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 23 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:39:

CCTCTTCAGT GTCTTCCTAT TGA   23

(2) INFORMATION FOR SEQ ID NO:40:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:40:

GGGAGGAGGT TGTAGGCAAC   20

(2) INFORMATION FOR SEQ ID NO:41:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 21 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:41:

AGCAAAGCAA AGGTGGCACA C   21

(2) INFORMATION FOR SEQ ID NO:42:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 24 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:42:

TGACATGGGA GAAGACACAC TTCC   24

(2) INFORMATION FOR SEQ ID NO:43:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 20 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:43:

AGGTTTACCA ATGTGTTTGG   20

(2) INFORMATION FOR SEQ ID NO:44:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 21 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:44:

TCTACATCCC ATTCTCTTCT G   21

(2) INFORMATION FOR SEQ ID NO:45:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 939 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..939
        (D) OTHER INFORMATION: /note= "putative human cyclophilin gene
                                from genomic clone (BAC clone 97) with
                                homology to rat cyclophilin cDNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:45:

```
TGTGATATTG ATTCATGCCC TCTTGCACCT TGCCAAACAT CACACGCTTG CCATCCAGTC      60

CACTCGATTT TGGCAGTGCA GATGAAAAAC TGGGAACCAT TTGTGTTGAG TCCAGCAAGA     120

TGCCAGGACC TGCATGTTTC AGAACGAAGT TCTTCATCAT CCAATTTCTC CCTGTATATG     180

GGCTTACCAC NACTGCCGTT AAGTCGTGTN AAGTCACCAC TCAGGTACAT AATGGAATAA     240

TTCTGCAAAG GCAGGAGNCA CTTTCTCTCC AGTGCTCAGA CCATGAAAGT TTTCTGATGT     300

CTTTGGAACT TTGTCTGCAA ATAGCTCGAA GGAGACATGG CCTAAAGGCT CGCCATCTGC     360

GGTGATATTG NAACATGGTA GGGCTGACCG TGGCTGTGGC CATGACTTTT TAGANTNNNN     420

NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN     480

NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNCCCAAT GCGGGACAGA GAATCNAAGA     540

AACTGTATTA GGGAAAGGGT CCTGAGTTTA TGCCAAAGTT TCCCAGATTG GTTTCCATTG     600

AAACGTAGCT CTGTGAGATA CCATCAGGTG TTATGTGAAG AAATGTCTGT GTAGTCAAAT     660

ATGTTTGAGT GAGTGAGCCT GAGCTGAGCA AGACTTTACT GCAAGACTTC CCATCTTCTG     720

TCCCTTTTTA TGCTAATGGG TAACACAAAC TCCAAAAGTG GGGTGTACAG CATGAGGCAT     780

TAACAAAAAT TTATTGGACC CCACACACNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN     840

NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN     900

NNNNNNNNNN NNNNNNNNNN NNNNNNNNNN NNNNCTCTC                            939
```

(2) INFORMATION FOR SEQ ID NO:46:

    (i) SEQUENCE CHARACTERISTICS:
      (A) LENGTH: 112 base pairs
      (B) TYPE: nucleic acid
      (C) STRANDEDNESS: single
      (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (ix) FEATURE:
      (A) NAME/KEY: -
      (B) LOCATION: 1..112
      (D) OTHER INFORMATION: /note= "positions 261-372 of putative
         human cyclophilin gene genomic clone
         (BAC clone 97) complementary strand with
         homology to positions 64-175 of rat
         cyclophilin gene cDNA"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:46:

```
GAAAGAGAGG TCACGAGTCT GGTACTTTCA AAAGACTACA GAAACCTTGA AACAGACGTT      60

TATCGAGCTT CCTCTGTACC GGATTTCCGA GCGGTAGACG CCACTATAAC NT            112
```

(2) INFORMATION FOR SEQ ID NO:47:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 112 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..112
        (D) OTHER INFORMATION: /note= "positions 64-175 of rat
                            cyclophilin cDNA with homology to
                            positions 261-372 of putative human
                            cyclophilin gene genomic clone (BAC
                            clone 97) complementary strand"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:47:

TTCGACATCA CGGCTGATGG CGAGCCCTTG GGTCGCGTCT GCTTCGAGCT GTTTGCAGAC   60

AAAGTTCCAA AGACAGCAGA AAACTTTCGT GCTCTGAGCA CTGGGGAGAA AG   112

(2) INFORMATION FOR SEQ ID NO:48:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 58 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..58
        (D) OTHER INFORMATION: /note= "positions 60-117 of putative
                            human cyclophilin gene genomic clone
                            (BAC clone 97) complementary strand
                            with homology to positions 348-405 of
                            rat cyclophilin gene cDNA"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:48:

GGTGAGCTAA AACCGTCACG TCTACTTTTT GACCCTTGGT AAACACAACT CAGGTCGT   58

(2) INFORMATION FOR SEQ ID NO:49:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 58 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

(ix) FEATURE:
    (A) NAME/KEY: -
    (B) LOCATION: 1..58
    (D) OTHER INFORMATION: /note= "positions 348-405 of rat
                              cyclophilin gene cDNA with homology to
                              positions 60-117 of putative human
                              cyclophilin gene genomic clone (BAC
                              clone 97) complementary strand"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:49:

TGCTGGACCA AACACAAATG GTTCCCAGTT TTTTATCTGC ACTGCCAAGA CTGAGTGG   58

(2) INFORMATION FOR SEQ ID NO:50:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 48 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..48
        (D) OTHER INFORMATION: /note= "positions 13-60 of putative
                                human cyclophilin gene genomic clone
                                (BAC clone 97) complementary strand with
                                homology to positions 404-451 of rat
                                cyclophilin gene cDNA"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:50:

AGTACGGGAG AACGTGGAAC GGTTTGTAGT GTGCGAACGG TAGGTCAG   48

(2) INFORMATION FOR SEQ ID NO:51:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 48 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..48
        (D) OTHER INFORMATION: /note= "positions 404-451 of rat
                                cyclophilin gene cDNA with homology to
                                positions 13-60 of putative human
                                cyclophilin gene genomic clone (BAC
                                clone 97) complementary strand"

(xi) SEQUENCE DESCRIPTION: SEQ ID NO:51:

GGCTGGATGG CAAGCATGTG GTCTTTGGGA AGGTGAAAGA AGGCATGA   48

(2) INFORMATION FOR SEQ ID NO:52:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 38 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..38
        (D) OTHER INFORMATION: /note= "positions 116-153 of putative human cyclophilin gene genomic clone (BAC clone 97) complementary strand with homology to positions 299-336 of rat cyclophilin gene cDNA"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:52:

GTTCTACGGT CCTGGACGTA CAAAGTCTTG CTTCAAGA   38

(2) INFORMATION FOR SEQ ID NO:53:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 38 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

    (ix) FEATURE:
        (A) NAME/KEY: -
        (B) LOCATION: 1..38
        (D) OTHER INFORMATION: /note= "positions 299-336 of rat cyclophilin gene cDNA with homology to positions 116-153 of putative human cyclophilin gene genomic clone (BAC clone 97) complementary strand"

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:53:

AGAACTTCAT CCTGAAGCAT ACAGGTCCTG GCATCTTG   38

(2) INFORMATION FOR SEQ ID NO:54:

    (i) SEQUENCE CHARACTERISTICS:
        (A) LENGTH: 28 base pairs
        (B) TYPE: nucleic acid
        (C) STRANDEDNESS: single
        (D) TOPOLOGY: linear

    (ii) MOLECULE TYPE: DNA (genomic)

```
(ix) FEATURE:
     (A) NAME/KEY: -
     (B) LOCATION: 1..28
     (D) OTHER INFORMATION: /note= "positions 229-256 of putative
                            human cyclophilin gene genomic clone
                            (BAC clone 97) complementary strand with
                            homology to positions 193-220 of rat
                            cyclophilin gene cDNA"
```

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:54:

TATTACCTTA TTAAGACGTT TCCGTCCT   28

(2) INFORMATION FOR SEQ ID NO:55:

```
     (i) SEQUENCE CHARACTERISTICS:
         (A) LENGTH: 28 base pairs
         (B) TYPE: nucleic acid
         (C) STRANDEDNESS: single
         (D) TOPOLOGY: linear

     (ii) MOLECULE TYPE: DNA (genomic)

     (ix) FEATURE:
          (A) NAME/KEY: -
          (B) LOCATION: 1..28
          (D) OTHER INFORMATION: /note= "positions 193-220 of rat
                                 cyclophilin gene cDNA with homology to
                                 positions 229-256 of putative human
                                 cyclophilin gene genomic clone (BAC
                                 clone 97) complementary strand"
```

    (xi) SEQUENCE DESCRIPTION: SEQ ID NO:55:

TCCTCCTTTC ACAGAATTAT TCCAGGAT ·   28